Lösungen zu den Übungsaufgaben in Kapitel 12

(1) In den Online-Materialien zu diesem Buch finden Sie die Datei »daten_kap11.txt«. Mit diesen Daten hatten Sie bereits in Kapitel 11 gearbeitet. Stellen Sie sich vor, es handele sich nicht – wie in Übungsaufgabe 1 in Kapitel 11 beschrieben – um ein Zwei-Gruppen-Experiment, bei dem die Versuchspersonen randomisiert einer von zwei experimentellen Bedingungen zugewiesen wurden, sondern vielmehr um ein Experiment mit intraindividueller Bedingungsvariation (Messwiederholung). Die insgesamt 40 Versuchspersonen werden zunächst in gute Stimmung versetzt, anschließend sollen sie eine Anagrammaufgabe lösen. Dann werden sie in schlechte Stimmung versetzt und sollen schließlich eine zweite Anagrammaufgabe lösen. Die Ergebnisse des ersten Messzeitpunkts (gute Stimmung) stehen in der ersten Spalte, die Ergebnisse des zweiten Messzeitpunkts (schlechte Stimmung) stehen in der zweiten Spalte.
Lesen Sie diese Datendatei in ein Statistikprogramm (z. B. R oder SPSS) oder ein Tabellenkalkulationsprogramm (z. B. Excel) ein. Beantworten Sie die folgenden Fragen:

(a) Prüfen Sie mit einem *t*-Test für abhängige Stichproben, ob die mittlere Differenz zwischen den beiden Messzeitpunkten auf dem 5 %-Niveau (einseitig) statistisch bedeutsam ist.

Die mittlere Differenz beträgt $\bar{x}_D = 2{,}25$. Der Standardfehler der mittleren Differenz beträgt $\hat{\sigma}_{\bar{x}_D} = \frac{\hat{\sigma}_D}{\sqrt{n}} = \frac{5{,}29}{\sqrt{40}} = 0{,}84.$ Der Wert der Prüfgröße beträgt $t = \frac{2{,}25}{0{,}84} = 2{,}69.$ Der kritische Wert unter der Nullhypothese bei α = 5% (einseitig) beträgt bei df = 39 Freiheitsgraden $t_{(0{,}95;39)}$ = 1,685. Der empirische *t*-Wert liegt darüber. Die Nullhypothese, der zufolge sich die kognitive Leistungsfähigkeit in Abhängigkeit von der Stimmungsinduktion nicht ändert oder sogar besser wird, kann daher verworfen werden.

(b) Wie lautet die empirische Effektstärke d''_2? Bewerten Sie die Größe von d''_2 anhand der Taxonomie von Cohen.

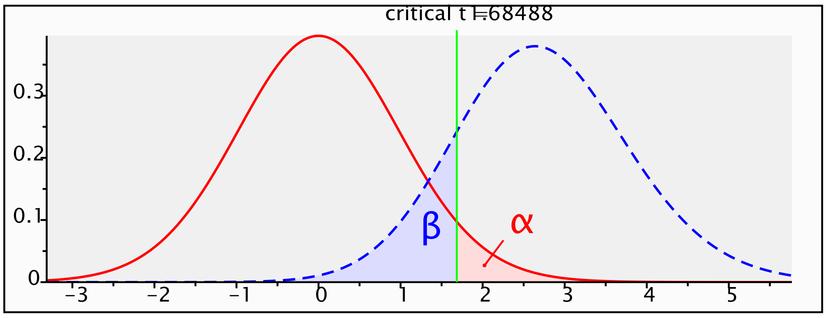
Gemäß Formel F 12.9 ergibt sich ein Wert von $d''_2 = \frac{\bar{x}_D}{\hat{\sigma}_D} = \frac{2{,}25}{5{,}29} = 0{,}425$. Es handelt sich um einen mittleren bis großen Effekt.

(c) Wo liegen die Grenzen des zweiseitigen 95 %-Konfidenzintervalls für die Effektgröße δ''?

Die Nonzentralitätsparameter der beiden gesuchten nonzentralen *t*-Verteilungen bestimmen wir mit Hilfe des Programms NDC. Bei t = 2,69 und df = 39 erhalten wir die Werte λ_{u} = 0,626 und λ_{o} = 4,72. Umgerechnet in δ'' (siehe Formel F 12.11) ergeben sich folgende Werte für die Grenzen des zweiseitigen 95 %-Konfidenzintervalls: δ''_{u} = 0,10 und δ''_{o} = 0,75.

(d) Wie groß war die Wahrscheinlichkeit, einen Effekt der gefundenen Größe (d''_2) oder einen noch größeren bei einem Signifikanzniveau von 5 % zu finden, falls der Effekt wirklich existiert?

Mit Hilfe des Programms G*Power (Testfamilie: t-Tests; Statistischer Test: Differenz zwischen zwei abhängigen Mittelwerten; Poweranalysetyp: Post hoc) ermitteln wir beim einseitigen Test, Effektgröße $d''_2 = 0{,}425$, α = 0,05 und n = 40 einen Wert von $1 - \beta$ = 0,84.



(e) Führen Sie mit diesen Daten einen Vorzeichentest durch. Prüfen Sie die Signifikanz des Ergebnisses auf der Basis eines *z*-Tests (α = 5 %, einseitig).

Es gibt 32 Paare mit positivem Vorzeichen und 8 mit negativem Vorzeichen der Messwertedifferenz ($x_{m1} - x_{m2}$). Da die Prüfgröße (S) der absoluten Häufigkeit der Paare mit positivem Vorzeichen entspricht, ergibt sich $s = 32$. Umgerechnet in einen z-Wert (siehe Formel F 12.13) ergibt sich $z = \sqrt{n} \cdot \frac{s/n - 0{,}5}{0{,}5} = \sqrt{40} \cdot \frac{32/40 - 0{,}5}{0{,}5} = 3{,}795$. Der kritische Wert unter der Nullhypothese bei $\alpha = 5\%$ (einseitiger Test) beträgt $z_{krit} = 1{,}645$. Der empirische z-Wert liegt darüber. Die Nullhypothese kann also verworfen werden.

(2) Bleiben wir bei dem in Aufgabe 1 geschilderten Experiment, in dem die Versuchspersonen einmal in gute und einmal in schlechte Stimmung versetzt wurden. Es soll nun geprüft werden, ob die Stimmung auch die Bereitschaft zur Hilfeleistung beeinflusst. Die Versuchspersonen sollten sich nach jeder der beiden Stimmungsinduktionen entscheiden, ob sie die 10 Euro, die sie als Entlohnung für ihre Teilnahme an der Untersuchung erhalten haben, (a) selbst behalten oder (b) einem Kinderhilfswerk spenden. Die Ergebnisse sind in der nachfolgenden Tabelle zusammengefasst.

		Schlechte Stimmung (Zeitpunkt 2)		Summe:
		y_1: selbst behalten	y_2: spenden	
Gute Stimmung (Zeit-punkt	x_1: selbst behalten	$n_{11} = 7$	$n_{12} = 2$	$n_{1\bullet} = 9$
	x_1: spenden	$n_{21} = 12$	$n_{22} = 19$	$n_{1\bullet} = 31$
Summe:		$n_{\bullet 1} = 19$	$n_{\bullet 2} = 21$	$n = 40$

(a) Prüfen Sie mit Hilfe eines McNemar-Tests auf der Basis der χ^2-Verteilung, ob es mehr Wechsler von »selbst behalten« nach »spenden« als von »spenden« nach »selbst behalten« gibt (α = 5 %, gerichtete Hypothese). Zu welchem Ergebnis kommen Sie, und was bedeutet dieses Ergebnis inhaltlich?

Das statistische Hypothesenpaar lautet:

H_0: $\pi_{12} \geq \pi_{21}$

H_1: $\pi_{12} < \pi_{21}$

Gemäß Formel F 12.17 erhalten wir für den Wert der Prüfgröße $\chi^2 = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}} = \frac{(2-12)^2}{2+12} = 7{,}143$.

Der kritische Wert lautet $\chi^2_{(0{,}95;1)} = 3{,}841$. Der empirische Wert liegt darüber. Die Nullhypothese, der zufolge es genau so viele Wechsler von »selbst behalten« nach »spenden« als von »spenden« nach »selbst behalten« gibt, kann also verworfen werden. Das bedeutet inhaltlich: Nachdem sie von guter in schlechte Stimmung versetzt worden sind, ist es wahrscheinlicher, dass man sich umentscheidet, wenn man sich vorher dazu entschieden hatte, das Versuchspersonengeld zu spenden. Schlechte Stimmung führt offenbar dazu, dass man sich eigennütziger verhält.

(b) Wie lautet die empirische Effektstärke, ausgedrückt als Odds-Ratio? Interpretieren Sie diesen Wert inhaltlich.

Das Odds-Ratio beträgt in unserem Beispiel $OR = n_{12}/n_{21} = 2/12 = 0{,}167$. Der Kehrwert, der hier besser interpretierbar ist, beträgt $n_{21}/n_{12} = 12/2 = 6$. Die Chance, von »spenden« nach »selbst behalten« zu wechseln, ist also um das Sechsfache größer als umgekehrt.

(c) Wie viele Personen wären in diesem Experiment benötigt worden, um einen Effekt der in Aufgabe (b) ermittelten Größe (OR) mit einer Wahrscheinlichkeit von 95 % zu finden (α = 5 %; gerichtete Hypothese), falls ein solcher Effekt in der Population tatsächlich existiert (Post-hoc-Poweranalyse)? Führen Sie mit G*Power eine Stichprobenumfangsplanung für den exakten Binomialtest durch.

In G*Power (Testfamilie: Exakt; Statistischer Test: Verhältnisse, zwei abhängige Gruppen [McNemar]; Poweranalysetyp: A priori) geben wir die folgenden Werte ein: Einseitiger Test; $OR_{\text{pop}} = 0{,}167$; $\alpha = 0{,}05$; $1 - \beta = 0{,}95$; Wechselwahrscheinlichkeit (Proportion of discordant pairs; siehe Formel F 12.25; hier geschätzt aus der empirischen relativen Häufigkeit der Wechsler): $\pi_{W1} = \pi^*_{12} + \pi^*_{21} = 2/40 + 12/40 = 0{,}35$. Dann erhalten wir eine optimale Stichprobengröße von $n = 52$.